

Judith Browne-Dianis
Linda Darling-Hammond
Carl Glickman
John Goodlad
Gloria Ladson-Billings
Deborah Meier
Larry Myatt

Pedro Noguera
Wendy Puriefoy
Sharon Robinson
Ted and NancySizer
Angela Valenzuela
George Wood,
Executive Director

Beyond Standardized Tests: Investing in a Culture of Learning

Introduction

Inspiration, hunger: these are the qualities that drive good schools. The best we educational planners can do is to create the most likely conditions for them to flourish, and then get out of their way.

— Ted Sizer¹

The late Ted Sizer, a life-long progressive educator and public school advocate, believed that the ultimate purpose of schooling is to teach students to “learn to use their mind well.”

Tragically, Ted’s simply stated, universally agreed upon purpose of schooling is not sufficiently reflected in the most popular current education reform strategies. Indeed, as long as we continue to tolerate the illusion that student learning can be sufficiently determined by a single day’s performance on a standardized test, we will continue to hinder educators’ efforts to create the optimal conditions for healthy, high-functioning learning environments that prepare all children to be life-long learners as well as productive and engaged citizens in our democracy.

When the patterns of our modern education system first took root more than a century ago, the objective was to process large numbers of people for the basic skills needed in low and unskilled jobs. This, as many educators of the time noted, was not structured to engage children in developing the skills and dispositions democratic life requires. Unfortunately, most of the contours of this system are still in place — a factory style model of schooling where young people are batched processed with a primary focus on low-level skills that can be measured on standardized examination. To our peril, we continue to focus on the rote learning of discrete bits of information with little or no attention paid to the higher order thinking skills our democracy, as well as our economy, requires.

As all other high-achieving nations have recognized, schools must focus much more on the ability to find, analyze, and use information in new ways than to remember large numbers of discrete facts. Indeed, in the four years from 1999 to 2003, the amount of new information produced approximately equaled the amount produced in the entire history of the world previously.² The transmission curriculum that dominated schools for the last 100 years — which assumed a stable body

of knowledge that could be codified in textbooks and passed onto students who would “learn” it by remembering bits of information — is counterproductive today, as are tests that focus largely on recall and recognition, rather than on production and application of what students have learned.

We can do better.

The U.S. education system began to adjust to these new demands in the early 1990s, with the development of new standards for learning stimulated by the Goals 2000 Act. However, with political pressures, many of these standards documents have evolved to suggest a curriculum that is a mile wide and an inch deep — one that requires teachers to rush through too many topics superficially, rather than to teach core ideas deeply.

By contrast, most high-achieving countries teach (and test) fewer topics each year and teach them more thoroughly so students build a stronger foundation in their learning.³ Their assessments require students to conduct research and scientific investigations, solve complex real-world problems in mathematics, and defend their ideas orally and in writing. This focuses students’ and teachers’ attention on the skills that democracy, higher education, and the jobs of tomorrow will require. Teachers’ involvement in scoring these assessments also helps them understand the standards deeply and develop curricula, with their colleagues, to support stronger learning — making the assessments more valuable for the school improvement process.

Unfortunately, many U.S. states that began to develop performance assessments like these in the 1990s have had to abandon them for traditional multiple-choice tests that more easily satisfy the requirements of No Child Left Behind. These tests often reinforce a curriculum that promotes rote learning rather than one that nurtures deep understanding. As tests have been used for accountability purposes, they have increasingly shaped teaching, reducing the amount of time students spend reading real books, writing essays, and conducting research. Indeed, scores of teachers worry that preparing for this kind of test does not prepare children for real-world applications of knowledge. As one Texas teacher noted in a survey, echoing the views of a large majority:

I have seen more students who can pass the [state test] but cannot apply those skills to anything if it’s not in the test format. I have students who can do

the test but can't look up words in a dictionary and understand the different meanings.... As for higher quality teaching, I'm not sure I would call it that. Because of the pressure for passing scores, more and more time is spent practicing the test and putting everything in the test format.⁴

Predictably, research on high-stakes accountability systems show that “what is tested is what is taught,” and that standards not represented on high stakes tests receive short shrift in the curriculum.⁵ A recent national survey of teachers found that teachers in high-stakes testing states were more likely than those in other states to report that they feel pressured to teach in ways that contradict their ideas of sound instructional practice. Teachers in these states reported not only that they assign less research and writing work, but that they cannot use computers to teach writing because the state test is handwritten.⁶

In addition, another recent study by the GAO laid bare the concerns of a growing number of researchers, who worry that “current accountability systems’ focus on test scores and mandated timelines for achieving proficiency levels for students do not give teachers enough time to learn, practice, and reflect on instructional practices and may discourage some teachers from trying ambitious teaching practices envisioned by standards-based reform.”⁷

Other studies have demonstrated that students are less likely to engage in extended research, writing, complex problem solving, and experimentation when the accountability system emphasizes short-answer responses to formulaic problems. These higher order modes of thinking are those very skills that often are cited as essential to maintaining America’s competitive edge and necessary for succeeding on the job, in college, and in life. As described by Achieve, a national organization of governors, business leaders, and education leaders, the problem with measures of traditional on-demand tests is that they cannot measure many of the skills that matter most for success in the worlds of work and higher education:

States ... will need to move beyond large-scale assessments because, as critical as they are, they cannot measure everything that matters in a young person's education. The ability to make effective oral arguments and conduct significant research projects are considered essential skills by both employers and postsecondary educators, but these skills are very difficult to assess on a paper-and-pencil test.⁸

If students are to be expected to achieve at higher standards, it stands to reason that educators must meet higher standards as well. They must know how to teach in ways that enable students to master challenging content and develop higher order thinking and performance skills, and that address the specific needs of different learners, including new English language learners as well as students with special education needs.

Clearly, there is a pressing need to rethink how we invest in our system of public education if we are to equip

our young people to meet the demands of citizenship and productivity in a global 21st century society. With the impending reauthorization of the Elementary and Secondary Education Act (ESEA), the time is ripe to restore the focus of education reform to its rightful bull’s-eye — on *learning*, and the core conditions that best support it. We must invest in the creation of a long-term teaching profession, not a short-term teaching force. And we must ensure a fair and equitable distribution of resources for education in all communities.⁹

Supporting Engaging and Challenging Learning Environments in All Our Public Schools

The real test for a student is when she is presented with something that is unfamiliar and asked to use what she has learned and the habits of hard thinking to make sense of it. That is, to make the unfamiliar familiar.

— Ted Sizer¹⁰

It is critically important to focus the learning enterprise on the right goals. This should include encouraging thoughtful measures of student performance that can support the kind of learning we need in schools, and developing a more useful method for charting school progress. Additionally, teachers should be supported in developing and using these measures in ways that inform instruction and curriculum. Although No Child Left Behind calls for multiple measures and for assessing higher-order thinking skills, it currently lacks incentives to encourage better assessments. This both undermines instructional quality and reinforces inequality, because low-income schools are most likely to experience an impoverished curriculum¹¹ organized primarily around narrow, lower-level tests.

To address these problems, the federal government should:

Support the development and use of assessments that measure the knowledge and skills needed in the 21st century.

1

In high-achieving countries, assessments routinely include evidence of actual student performance on challenging tasks that evaluate standards of advanced learning. Curricula and assessments emphasize deep knowledge of core concepts within and across the disciplines, problem solving, analysis, synthesis, and critical thinking. As a large and increasing part of their examination systems, high-achieving nations use open-ended performance tasks along with curriculum-embedded assessments to give students opportunities to develop and demonstrate 21st century abilities: the abilities to find and organize information to solve problems, frame and conduct inquiries, analyze and synthesize data, and apply learning to new situations. The curricular and assessment systems evaluate students’ abilities in research papers, investigations and experiments, open-ended tasks, and oral presentations that are developed and scored to common specifications,

as well as examinations that include essays and open-ended problems, usually given at the end of a course or year. For comparable practices to be undertaken in the United States, the government should:

- ***Fund an intensive development effort that enables states or consortia of states to work with assessment experts to develop, validate, and test high-quality performance assessments and to train the field of practitioners*** — ranging from a new generation of state and local curricular and assessment specialists to teachers and leaders — who can be involved in the development, administration, and scoring of these assessments in valid and reliable ways. The federal government should also ***fund high quality research on the validity, reliability, instructional consequences, and equity consequences of these assessments.***
- ***Encourage improvements in state and local assessment practices.*** To model high-quality items and better measure standards of learning, the federal government should move the National Assessment of Educational Practice toward a more performance-oriented assessment, as it was when it was first launched in the 1950s, with tasks that evaluate students' abilities to solve problems, explain, and defend their ideas. The new Elementary and Secondary Education Act should provide incentives and funding for states or state consortia to refine their state assessments, and introduce related, high-quality performance assessments that evaluate critical thinking and applied skills. It should also support states in making such assessments reliable, valid, and practically feasible, through scorer training and moderation systems.
- ***Ensure more appropriate assessment for special education students and English language learners (ELLs) by underwriting efforts to develop, validate, and disseminate more appropriate assessments*** in the content areas for these students, and by ensuring that the law and regulations encourage assessments that are based on professional testing standards for these groups. This would include helping to develop and requiring the use of tests that are language-accessible for ELLs and appropriate for special education students, and evaluating their gains at all points along the achievement continuum.

Periodic high-quality external assessments should include open-ended items, such as essays and problem solutions that call on students to evaluate and analyze information, and communicate and defend their ideas, along with selected-response items that focus more intently on analysis. These reference examinations do not need to be conducted on an every child, every year basis to be useful for curricular planning and school monitoring, and may be of higher quality if they are less frequent or conducted on a sampling basis, as is true in other nations. Most countries assess students in one or two grade levels before high school, and then create a more extensive system of college and career-ready assessments, developed by high school and university faculty together, to evaluate older students' abilities. Both sets of assessments often include syllabus-based within-classroom evaluations of student work that are

Models for Measuring Student Learning: Mission Hill School

The Mission Hill School is a Boston Public School, serving children in grades K-8. A small community, with approximately 170 students, Mission Hill's classrooms consist of no more than 20 students, and are multi-age; most children spend two years with the same teacher. About half of the school day is set aside for thematic work, which may include reading and writing, research, art and construction, and practical mathematic applications. In addition, there is time set aside for reading, writing and mathematics — which may or may not relate to the thematic work. All instruction is integrated so that reading, writing and math are taught each day within the thematic structure.¹³ Mission Hill has developed a full K-8th grade approach to assessing students that is semi-numerical, narrative, and descriptive by categories: semi-annual reading interviews are taped and "scored", as are writing samples that include one "standardized" piece and one chosen by the teacher (one fiction and one nonfiction) — scored by the faculty, and other work samples. In 5th grade, students review their own portfolios and compose a description of what they have learned about themselves at Mission Hill. In addition, they join a graduation committee, as an assistant to a 7th grader, whom they follow until they themselves begin to prepare for their own graduation exercises (which start in mid-7th grade and are completed by spring of 8th grade). In 6th grade, students are assessed in geography. After children are fluent readers on grade-level material, teachers use the school's protocols for discussions around literature and students keep a running record of their reading. 8th grade students study the effects of media in society and prepare for their portfolio requirements necessary for graduation. There are five graduation exams: one each in literature/writing, science, math, history and the arts in the spring of a student's 8th grade year.⁹

part of the examination scores. At the high school level, ambitious intellectual tasks include research papers and presentations, literary analyses, science experiments, complex mathematical problem solutions and models, uses of technology, and exhibitions of learning from the arts. This approach creates much more equitable curricular opportunities, greater attention to student learning, and comparability in evaluation, because of investments in teacher training and moderation.

These kinds of systems, if accompanied by adequate investments in teacher knowledge and skill, could both improve the overall quality of instruction and learning in American schools and leverage gains for students who will otherwise experience a low-level curriculum further impoverished by non-diagnostic teaching. Efforts to create a "thinking curriculum" for all students are critically important to individual futures and our national

welfare. Such efforts are unlikely to pay off, however, unless other central changes are made in the ways tests are used and accountability systems are designed, so that new standards and assessments inform more skillful and adaptive teaching that enables more successful learning for all students.

2 Support states to develop a diagnostic School Quality Review system to evaluate schools, to guide improvement, and to share best practices.

I and my fictitious friend Horace Smith identified “five imperatives for better schools.” We wrote in Horace’s Compromise:

- *Give room to teachers and students to work in their own, appropriate ways.*
- *Insist that students clearly exhibit the mastery of their schoolwork.*
- *Get the incentives right, for students and for teachers.*
- *Focus on the students’ use of their minds.*
- *Keep the structure simple and thus flexible.*

— Ted Sizer¹⁶

Having data on school outcomes is not enough to guide meaningful improvement. An expert analysis of school capacity and needs is a necessary first step toward identifying problems and moving strategically toward solutions. School-level improvement can be supported by school inspection systems, like those common in many other nations, in which trained experts, usually highly-respected former practitioners, evaluate schools by spending several days visiting classrooms, examining samples of student work, and interviewing students about their understanding and their experiences, as well as looking at objective data such as test scores, graduation rates, and attendance and disciplinary rates. In some cases, principals accompany the inspectors into classrooms and are asked for their own evaluations of the lessons. In this way, the inspectors are able to make judgments about the instructional and supervisory competence of principals. Inspectors may also play a role in assuring the quality and comparability of school-based assessments (as in England and Australia), as well as schools’ internal assessment and evaluation processes (as in Hong Kong).

In most countries’ inspection systems, schools are rated on the quality of instruction and other services and supports, as well as students’ performance and progress on a wide range of dimensions. Schools receive extensive feedback on what the inspectors both saw and recommend. Reports are publicly posted. Schools requiring intervention are then given more expert attention and support, and placed on a more frequent schedule of visits. Those that persistently fail to pass may be placed under local government control and could be closed down if they are not improved. Successful schools are also highlighted and studied.

Models for Measuring Student Learning: Federal Hocking High School

Federal Hocking High School, a small school (grades 7–12) set in rural Ohio, focuses its efforts on three primary goals: 1) Developing a love of learning within each student regardless of his or her background, 2) Inspiring active citizenship among its students, and 3) Giving every student the option to be and do anything they choose upon graduation. Each student at Federal Hocking finishes his or her high school career with the opportunity to leave a lasting impact on the school and/or community through the Senior Project experience. The Senior Project gives students the chance to gain new skills, see how they can make a difference in the community, and pursue, in-depth, one of their interests. In addition, all students complete a Graduation Portfolio, which is designed to represent what students valued in their high school experience. There are two goals for the Graduation Portfolio: 1) to enable the faculty of FHHS to review each graduate’s readiness to enter the world after high school, and 2) to enable each student to reflect on his or her education and how prepared he or she is for the responsibilities of democratic citizenship, the world of work, and a life of learning.¹⁴

In June 2009, the Broader, Bolder Approach (BBA) Accountability Committee proposed the introduction of a U.S. Inspectorate system, stating that quantitative measures of a school’s performance offer only a partial picture of school quality. BBA offers the following example:

...If we want to know whether test score gains have been produced by an undue emphasis on test preparation and low-level instruction focused exclusively on factual recall, qualified evaluators have to visit classrooms, observe instruction, review teacher lesson plans and student assignments, and look at samples of student work. Only by visiting schools and classrooms can inspectors assess whether a school maintains a safe and orderly environment; practices that are respectful of and engage students, staff, and parents; programs to promote student health and other non-cognitive outcomes; and a collegial professional culture in which teachers and administrators use all available data in a collaborative fashion to continuously improve the work of the school.¹⁷

An Americanized version of the inspectorate system, designed by former members of the British inspectorate with U.S. educators, has been piloted in several states and cities, including New York, California, Rhode Island, and Chicago. The process joins expert team leaders with current practitioners from other schools, and has proved an extremely effective strategy for enabling schools to get an objective look at their practices, creating an evidence base that complements outcome information,

and providing diagnostics and recommendations that are essential for any serious improvement to occur. When practicing educators are among the members of the teams, they also learn directly about colleagues' practices and how to evaluate education in ways that travel back with them to their own schools, creating a reciprocal and self-sustaining learning system across the state. The capacity for school reviews could be developed within state education agencies or through school accreditation agencies. Initially, it might be focused on the needs of schools requiring intervention. Ideally, it would be expanded to provide a means for continuous improvement for all schools across the state and as a coordinated means to share best practices.

If the federal government would build states' capacities to involve communities and to distribute timely, relevant, and descriptive information about student achievement, then powerful partnerships and coalitions supporting education could emerge. Just like the assessments for students, the aim of such information sharing would not be to punish or penalize schools and districts, but to inspire discussion and to coalesce a local community around striving toward world-class standards.

3

Organize regularly available, high-quality professional development around performance based assessment of higher order thinking skills and deep content knowledge.¹⁸

The federal government should support states to develop an infrastructure for high-quality professional development by funding professional development time and organizing the multiple resources of the state — from universities, districts, and nonprofit organizations — to ensure that expertise and capacity are developed to address the use of performance-based assessments and the teaching strategies that support all students, including ELLs and those children with special learning needs. As in a number of states, Teacher and Leadership Academies can play a key role in doing this, helping to organize intensive institutes and networks to support learning and teaching based on higher-order skills; training principals and teachers as mentors and coaches who can support others in districts; and providing materials and expertise to support professional learning at the school level. Districts should enable schools to tap both these external resources and the critical resource of expert teachers within the schools, by creating time and regular opportunities for peer coaching and collaborative learning.

There are a number of illustrative examples from elsewhere in the world about how to scale more reciprocal learning environments and practices. Since 2000, for example, the Australian government has been sponsoring the Quality Teacher Programme, a large-scale initiative that provides funding to update and improve teachers' skills and understandings in priority areas and enhance the status of teaching in both government and non-government schools. The Programme operates at three levels: (1) Teaching Australia (formerly the National Institute for Quality Teaching and School Leadership); (2)

Models for Measuring Student Learning: New York Performance Standards Consortium

The Consortium is a coalition of high schools across New York State that have pioneered the creation of educational communities synonymous with active student learning, exemplary professional development and innovative curricular and teaching strategies for 21st century students. Many of the participating schools and leaders have been nationally recognized through such honors as New American High Schools, 21st Century Schools, Blue Ribbon Schools of Excellence, and Compact for Learning Schools. Consortium schools have established a system of assessment that consists of eight components including alignment with state standards, professional development, external review, and formative and summative data. Consortium schools have documented how their work meets and exceeds New York State Regents standards through a system of rigorous commencement-level performance-based assessment tasks. Performance on these tasks is reflected on student transcripts and results are used for college admission. The tasks require students to demonstrate accomplishment in analytic thinking, reading comprehension, research writing skills, the application of mathematical computation and problem-solving skills, computer technology, the utilization of the scientific method in undertaking science research, appreciation of and performance skills in the arts, service learning and school to career skills. Experts external to the schools, from universities and the business world, participate in reviews of student work.¹⁵

National Projects; and (3) State and Territory Projects. Teaching Australia facilitates the development and implementation of nationally agreed upon teaching standards, conducts research and communicates research findings, and facilitates and coordinates professional development courses. The National Projects have a national focus and include programs designed to identify and promote best practice, support the development and dissemination of professional learning resources in priority areas, and develop professional networks for teachers and school leaders. The State and Territory Projects fund a wide variety of professional learning activities for teachers and school leaders under agreements with state and territory education authorities, allowing professional development activities to be tailored to local needs. These projects include school-based action research and learning, conferences, workshops, on-line or digital media, and training of trainers, school project and team leaders.

Similarly, Western Australia's highly successful Getting it Right (GiR) Strategy provides specialist teaching personnel, professional development, and support to select primary schools to improve literacy and numeracy outcomes of high needs students, with a focus on

Aboriginal and other at-risk students. Each school selects a highly regarded teacher with interest and expertise in numeracy or literacy to be a Specialist Teacher (ST), who is then trained through a series of seven three-day intensive workshops over the course of their initial two-year appointment. The Specialist Teachers work “shoulder to shoulder” with teachers in their schools, for about half a day each week for each teacher. The Specialist Teachers monitor and record student learning, help teachers analyze student learning, model teaching strategies, plan learning activities to meet the identified needs of students, assist with the implementation of these activities, and provide access to a range of resources, sharing expertise and encouraging teachers to be reflective about their practice. Teachers show greatly enhanced knowledge about how students’ learn reading, writing, and mathematics and much stronger teaching and assessment skills, including their ability to use data to identify and diagnose students’ learning needs and to plan explicit teaching approaches to address these needs.

Conclusion

It is a radical idea that all children grow at the same rate and in the same way and thus can thereby be accurately classified and “graded” in narrow, standardized ways. It is a radical idea that the power of a child’s mind can be plumbed by a single test and reduced to a small clutch of numbers. It is a radical idea that people of any age can learn well in crowded, noisy, and ill-equipped places. It is a radical idea that serious learning can best emerge from a student’s exposure to short blasts of “delivered” content, each of less than an hour in length, and unified by no coherent set of common ideas. It is a radical idea that a child can learn what is needed to live well in a complex society with schooling that encompasses barely half the days of a calendar year; and that ignores the opportunities — or lack of opportunities — available to each child.
— Ted Sizer¹⁹

Creating schools that enable all children to learn requires the development of systems that enable all educators and schools to become healthy, high-functioning democratic learning communities. Indeed, what is required is a capacity-building enterprise leveraged by clear, meaningful learning goals and intelligent, reciprocal accountability systems that guarantee skillful teaching in well-designed, adequately resourced schools for all learners. It is not only possible but imperative that America close the achievement gap among its children by addressing the yawning opportunity gap that denies these fundamental rights. Given the critical importance of education for individual and societal success in the flat world we now inhabit, inequality in the provision of education is an antiquated tradition the United States can no longer afford.

If No Child Left Behind is to be anything more than empty rhetoric, we will need a policy strategy that creates a 21st century curriculum for all students, and supports it with thoughtful assessments, access to knowledgeable, well-supported teachers, and equal access to school resources.

In sum, it is time to rethink learning. And now is the time to act.

Endnotes

- 1 Theodore R.Sizer, *Horace's Compromise: the Dilemma of the American School* (Houghton Mifflin: Boston, 1984).
- 2 Hal R Varian and Peter Lyman, "How Much Information?" (University of California at Berkeley, School of Information Management & Systems (SIMS), 2003), http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/printable_report.pdf (Accessed September 22, 2007)
- 3 Organisation for Economic Co-operation and Development, Programme for International Student Assessment 2006: Science Competencies for Tomorrow's World; William Schmidt, "What's Missing from Math Standards: Focus, Rigor and Coherence," *American Educator* (Spring 2008), http://www.aft.org/pubs-reports/american_educator/issues/spring2008/schmidt.htm (Accessed 4/1/08)
- 4 Joseph Pedulla, Lisa Abrams, George Madaus, Michael Russell, Miguel Ramos, & Jing Miao, Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers (Boston: National Board on Testing and Public Policy, Boston College, 2003); Walter Haney, "The Myth of the Texas Miracle in Education," *Educational Policy Analysis Archives* 8, no 41 (2000), <http://epaa.asu.edu/epaa/v8n41/> (Accessed March 28, 2008); Brett D Jones & Robert J Egly, "Voices from the Frontlines: Teachers' Perceptions of High-Stakes Testing," *Education Policy Analysis Archives* 12, no 39 (2004), <http://epaa.asu.edu/epaa/v12n39/> (Accessed August 10, 2004)
- 5 See for example, Walter Haney, "The Myth of the Texas Miracle in Education"; Joan L Herman & Shari Golan, Effects of Standardized Testing on Teaching and Schools," *Educational Measurement: Issues and Practice* 12, no 4 (1993): 20–25, 41–42; Brett D Jones & Robert J Egly, "Voices from the Frontlines"; M Gail Jones, Brett D Jones, Belinda Hardin, Lisa Chapman, & Tracie Yarbrough, "The Impact of High-Stakes Testing on Teachers and Students in North Carolina," *Phi Delta Kappan* 81, no 3 (1999): 199–203; Stephen P Klein, Laura S Hamilton, Daniel McCaffrey, & Brian M Stetcher, "What Do Test Scores in Texas Tell Us?" (Santa Monica: The RAND Corporation, 2000); Daniel M Koretz & Sheila I Barron, "The Validity of Gains on the Kentucky Instructional Results Information System (KIRIS)," (MR-1014-EDU, Santa Monica, CA: The RAND Corporation, 1998); Daniel M Koretz, Robert L Linn, Stephen B Dunbar, & Lorrie A Shepard (1991, April) The Effects of High-Stakes Testing: Preliminary Evidence about Generalization Across Tests, Paper presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, Chicago, IL ; Robert L Linn, "Assessments and Accountability," *Educational Researcher* 29, no 2 (2000): 4–16; Robert L Linn, M Elizabeth Graue, & Nancy M Sanders, "Comparing State and District Test Results to National Norms: The Validity of Claims that 'Everyone is Above Average '" *Educational Measurement: Issues and Practice* 9 (1990): 5–14; W James Popham, "Why Standardized Test Scores Don't Measure Educational Quality," *Educational Leadership* 56, no 6 (1999): 8–15; Mary Lee Smith, "Put to the Test: The Effects of External Testing on Teachers," *Educational Researcher* 20, no 5 (2001): 8–11.
- 6 Pedulla, et al , Perceived Effects of State-Mandated Testing Programs, 2003.
- 7 GAO, "Student Achievement: Schools Use Multiple Strategies to Help Students Meet Academic Standards, Especially Schools with Higher Proportions of Low-Income and Minority Students," (15–16, Washington, DC: US Government Accountability Office, 2009).
- 8 Achieve, "Do Graduation Tests Measure Up? A Closer Look at State High School Exit Exams," (Executive Summary, Washington, DC: Achieve, Inc , 2004).
- 9 To read the Forum's other policy recommendations regarding issues of teaching and fairness, visit forumforeducation.org or rethinklearningnow.com.
- 10 Theodore R. Sizer, speech given at CES Fall Forum, Providence, Rhode Island, 2000.
- 11 GAO, "Student Achievement: Schools Use Multiple Strategies to Help Students Meet Academic Standards, Especially Schools with Higher Proportions of Low-Income and Minority Students," (Washington, DC: US Government Accountability Office, 2009).
- 12 Theodore R. Sizer, speech given at CES Fall Forum, San Francisco, California, 1997.
- 13 Mission Hill School website: http://www.missionhillschool.org/mhs/Welcome_.html, (Accessed November 11, 2009).
- 14 Federal Hocking High School website: <http://www.federalhocking.k12.oh.us/extra/hs/index.htm> (Accessed November 11, 2009).
- 15 From the New York Performance Standards Consortium Website: <http://performanceassessment.org/> (Accessed November 11, 2009).
- 16 Theodore R. Sizer, speech given at CES Fall Forum, San Francisco, California, 1997.
- 17 Broader, Bolder Approach to Education Campaign, "School Accountability A Broader, Bolder Approach," (Pg. 5, Washington, DC, 2009).

- 18 For a more detailed look at innovative teacher policies, see the Forum's policy paper, "Effective Teachers, High Achievers: Investing in a Teaching Profession," available on forumforeducation.org or rethinklearningnow.com/principles/teaching/.
- 19 Theodore R.Sizer, speech given at CES Fall Forum, Washington, DC, 2002.